

# Unit 3

## Newton Forward And Backward Interpolation

**Interpolation** is the technique of estimating the value of a function for any intermediate value of the independent variable, while the process of computing the value of the function outside the given range is called **extrapolation**.

**Forward Differences** : The differences  $y_1 - y_0, y_2 - y_1, y_3 - y_2, \dots, y_n - y_{n-1}$  when denoted by  $\Delta y_0, \Delta y_1, \Delta y_2, \dots, \Delta y_{n-1}$  are respectively, called the first forward differences. Thus the first forward differences are :

Forward difference table

$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
$x_0$	$y_0$					
$x_1$ (= $x_0 + h$ )	$y_1$	$\Delta y_0$				
$x_2$ (= $x_0 + 2h$ )	$y_2$	$\Delta y_1$	$\Delta^2 y_0$			
$x_3$ (= $x_0 + 3h$ )	$y_3$	$\Delta y_2$	$\Delta^2 y_1$	$\Delta^3 y_0$		
$x_4$ (= $x_0 + 4h$ )	$y_4$	$\Delta y_3$	$\Delta^2 y_2$	$\Delta^3 y_1$	$\Delta^4 y_0$	
$x_5$ (= $x_0 + 5h$ )	$y_5$	$\Delta y_4$	$\Delta^2 y_3$	$\Delta^3 y_2$	$\Delta^4 y_1$	$\Delta^5 y_0$

### NEWTON'S GREGORY FORWARD INTERPOLATION FORMULA :

This formula is particularly useful for interpolating the values of  $f(x)$  near the beginning of the set of values given.  $h$  is called the interval of difference and  $u = (x - a) / h$ , Here  $a$  is first term.

$\theta^\circ$	$45^\circ$	$50^\circ$	$55^\circ$	$60^\circ$
$\sin \theta$	0.7071	0.7660	0.8192	0.8660

$x^\circ$	Differences			
	$10^1y$	$10^1\Delta y$	$10^1\Delta^2y$	$10^1\Delta^3y$
45°	7071	589		
50°	7660	532	-57	
55°	8192	468	-64	-7
60°	8660			

**Backward Differences :** The differences  $y_1 - y_0, y_2 - y_1, \dots, y_n - y_{n-1}$  when denoted by  $\nabla y_1, \nabla y_2, \dots, \nabla y_n$ , respectively, are called first backward difference. Thus the first backward differences are :

### Backward difference table

$x$	$y$	$\nabla y$	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$	$\nabla^5 y$
$x_0$	$y_0$					
$x_1$ (= $x_0 + h$ )	$y_1$	$\nabla y_1$				
$x_2$ (= $x_0 + 2h$ )	$y_2$	$\nabla y_2$	$\nabla^2 y_2$			
$x_3$ (= $x_0 + 3h$ )	$y_3$	$\nabla y_3$	$\nabla^2 y_3$	$\nabla^3 y_3$		
$x_4$ (= $x_0 + 4h$ )	$y_4$	$\nabla y_4$	$\nabla^2 y_4$	$\nabla^3 y_4$	$\nabla^4 y_4$	
$x_5$ (= $x_0 + 5h$ )	$y_5$	$\nabla y_5$	$\nabla^2 y_5$	$\nabla^3 y_5$	$\nabla^4 y_5$	$\nabla^5 y_5$

### NEWTON'S GREGORY BACKWARD INTERPOLATION FORMULA :

This formula is useful when the value of  $f(x)$  is required near the end of the table.  $h$  is called the interval of difference and  $u = (x - a_n) / h$ , Here  $a_n$  is last term.

#### Example :

Input : Population in 1925

<i>Year (x):</i>	1891	1901	1911	1921	1931
<i>Population (y): (in thousands)</i>	46	66	81	93	101

Output :

<i>x</i>	<i>y</i>	$V_y$	$V^2y$	$V^3y$	$V^4y$
1891	46	20			
1901	66	15	- 5		
1911	81	12	- 3	2	
1921	93	8	- 1	- 1	
1931	101				- 3

**Value in 1925 is 96.8368**

**Lagrange Interpolation Formula**

Lagrange’s formula is applicable to problems where the independent variable occurs at equal and unequal intervals, but preferably this formula is applied in a situation where there are unequal intervals for the given independent series. The values of the independent variables (X) are given as a,b,c,d,...a,b,c,d,... etc., and the corresponding values of the function (dependent variable) are f(a),f(b),f(c),f(d),...f(a),f(b),f(c),f(d),... etc. So the value of the function corresponding to a given value of X=Xo is given by f(Xo) which is calculated by the formula presented by Lagrange. Thus, if

XX	f(X)
aa	f(a)
bb	f(b)
cc	f(c)
dd	f(d)
::	::

f(Xo) is given as:

$$f(X_o) = \frac{(X_o - a)(X_o - b)(X_o - c)}{(a - b)(a - c)(a - d)} * f(a) + \frac{(X_o - a)(X_o - c)(X_o - d)}{(b - b)(b - c)(b - d)} * f(b) + \frac{(X_o - a)(X_o - b)(X_o - d)}{(c - a)(c - b)(c - d)} * f(c) + \frac{(X_o - a)(X_o - b)(X_o - c)}{(d - a)(d - b)(d - c)} * f(d) + \dots$$

Although this formula looks somewhat lengthy, is in fact very easy to remember and write. The first thing to note is that successive terms are multiplied

by  $f(a), f(b), f(c), \dots, f(a), f(b), f(c), \dots$  etc. respectively, and hence there will be as many terms as the number of items in the data. Thus, for example, if we have four yearly observations on the production of rice we will only need to write four such terms.

The numerator in each term contain factors

like  $(X_0 - a)(X_0 - b)(X_0 - c) \dots (X_0 - a)(X_0 - b)(X_0 - c) \dots$  etc. The point to note here is that the first term omits  $(X_0 - a)(X_0 - a)$ , the second term omits  $(X_0 - b)(X_0 - b)$ , the third term omits  $(X_0 - c)(X_0 - c)$  and so on.

The denominator in each term contains different factors of the independent variable. In the first term the difference is taken from  $a$ , in the second term the difference is taken from  $b$ , in the third term the difference is taken from  $c$  and so on.

When interpolating with this formula it is recommended that you do not look at the formula to write the values of  $a, b, c, \dots, a, b, c, \dots$  and  $f(a), f(b), f(c), \dots, f(a), f(b), f(c), \dots$  etc. but rather directly write down the values as if you are writing the formula with the given values. This is suggested because confusion arises while writing the values of  $a, b, c, a, b, c$  etc. in the formula

Interpolation refers to the process of creating new data points given within the given set of data. The above MATLAB code computes the desired data point within the given range of discrete data sets using the formula given by Gauss.

The gaussian interpolation comes under the Central Difference Interpolation Formulae which differs from Newton's Forward interpolation formula formula.

Suppose we are given the following value of  $y=f(x)$  for a set values of  $x$ :

X:  $x_0 \ x_1 \ x_2 \ \dots \ x_n$

Y:  $y_0 \ y_1 \ y_2 \ \dots \ y_n$

Then the process of finding the value of  $y$  corresponding to any value of  $x=x_i$  between  $x_0$  and  $x_n$  is called interpolation. Thus interpolation is the technique of estimating the value of a function for any intermediate value of the independent variable while the process of computing the value of the function outside the given range is called extrapolation. However, extrapolation is beyond the scope of this code.

Suppose that the function  $y=f(x)$  is tabulated for the equally spaced values  $x = x_0, x = x_0+h, x = x_0+2h, \dots, x = x_0+nh$ , giving  $y= y_0, y_1, y_2, \dots, y_n$ . To determine the value of  $f(x)$  or  $f'(x)$  for some intermediate values of  $x$ , the following three types of differences are found useful.

1. Forward Difference
2. Backward difference
3. Central Difference

The common Newton's forward formula belongs to the Forward difference category. However, the Gaussian forward formula formulated in the attached code belongs to the central difference method.

Gauss forward formula is derived from Newton's forward formula which is:

Newton's forward interpretation formula:

$$Y_p = y_0 + p \Delta y_0 + \frac{p(p-1)}{1 \cdot 2} \Delta^2 y_0 + \frac{p(p-1)(p-2)}{1 \cdot 2 \cdot 3} \Delta^3 y_0 + \dots$$

We have,

$$\Delta^2 y_0 - \Delta^2 y_{-1} = \Delta^3 y_{-1}$$

$$\text{i.e. } \Delta^2 y_0 = \Delta^2 y_{-1} + \Delta^3 y_{-1}$$

Similarly,

$$\Delta^3 y_0 = \Delta^3 y_{-1} + \Delta^4 y_{-1}$$

$$\Delta^4 y_0 = \Delta^4 y_{-1} + \Delta^5 y_{-1} \text{ etc..}$$

$$\text{Also, } \Delta^3 y_{-1} - \Delta^3 y_{-2} = \Delta^4 y_{-2}$$

$$\text{I.e. } \Delta^3 y_{-1} = \Delta^3 y_{-2} + \Delta^4 y_{-2}$$

Similarly  $\Delta^4 y_{-1} = \Delta^4 y_{-2} + \Delta^5 y_{-2}$  etc..

Substituting for  $\Delta^2 y_0$ ,  $\Delta^3 y_0$ ,  $\Delta^4 y_0$ ..... from (2),(3),(4)..... in (1), we get

$$Y_p = y_0 + p\Delta y_0 + \frac{p(p-1)(\Delta^2 y_{-1} + \Delta^3 y_{-1})}{1.2} + \frac{p(p-1)(p-2)(\Delta^3 y_{-1} + \Delta^4 y_{-1})}{1.2.3} + \frac{p(p-1)(p-2)(p-3)(\Delta^4 y_{-1} + \Delta^5 y_{-1})}{1.2.3.4} + \dots$$

$$\text{Hence } y_p = y_0 + p\Delta y_0 + \frac{p(p-1)\Delta^2 y_{-1}}{2!} + \frac{((p+1)p(p-1)\Delta^3 y_{-1})}{3!} + \frac{((p+1)p(p-1)(p-2)\Delta^4 y_{-2})}{4!} + \dots \text{ using (5).....}$$

Which is then called Gauss's forward interpolation formula.

One more representation of this formula is in the central difference notation, this formula will be

$$y_p = y_0 + p\delta y_{1/2} + \frac{p(p-1)\delta^2 y_0}{2!} + \frac{((p+1)p(p-1)\delta^3 y_{1/2})}{3!} + \frac{((p+1)p(p-1)(p-2)\delta^4 y_0)}{4!} + \dots$$

These interpolation formulae are applicable for interpolation near the beginning and end of tabulated values.

However, the Gaussian forward formula are best suited for interpolation near the middle of the table.

The coefficients in the central difference formula such as that of Gauss are smaller and converge faster than those in Newton's formulae. Therefore, as much as possible, central difference formulae should be used in preference to Newton's formulae.

The above code starts with asking the no. of data. The input data is denoted by x.

The difference between the two consecutive values of x is denoted by h.

One more parameter used in this formula is p, which is computed by the following formula

$$p = \frac{(x - x_0)}{h}$$

The corresponding value of y is also asked from the user.

After entering these values, the code computes the parameters h and p automatically and displays the result on calling.

The final desired answer is stored in a variable named ans and is displayed on calling in MATLAB window.

An example :-

No. of variables = 5

The x values are :- 21, 25, 29, 33, 37 and the corresponding values of y are given by

18.478 , 17.8144 ,  
17.8144 , 16.3432 and 15.5154.

$x_p$	$y_p$	$\Delta y_p$	$\Delta^2 y_p$	$\Delta^3 y_p$	$\Delta^4 y_p$	$\Delta^5 y_p$
21	-2	18.478	-0.6564	-0.0510	-0.0074	-0.0022
25	-1	17.8144	-0.7074	-0.0564	-0.0076	
29	0	17.8144	-0.7638	-0.0640		
33	1	16.3432	-0.8278			
37	2	15.5154				

On running the code for the desired value of  $y_{30}$ , we get the following values :-

$h=4$

$p=0.25$

Ans=16.9216

A formula in which the nodes (cf. [Node](#)) nearest to the interpolation point  $x$  are used as interpolation nodes. If  $x = x_0 + th$ , the formula

$$G_{2n+1}(x_0 + th) = f_0 + f_{1/2}^1 t + f_0^2 \frac{t(t-1)}{2!} + \dots +$$

$$+ f_0^{2n} \frac{t(t^2-1)\dots[t^2-(n-1)^2](t-n)}{(2n)!},$$

written with respect to the nodes  $x_0, x_0 + h, x_0 - h, \dots, x_0 + nh, x_0 - nh$  is called the Gauss forward interpolation formula, while the formula

$$G_{2n+1}(x_0 + th) = f_0 + f_{-1/2}^1 t + f_0^2 \frac{t(t+1)}{2!} + \dots +$$

$$+ f_0^{2n} \frac{t(t^2-1)\dots[t^2-(n-1)^2](t+n)}{(2n)!},$$

written with respect to the nodes  $x_0, x_0 - h, x_0 + h, \dots, x_0 - nh, x_0 + nh$  is called the Gauss backward interpolation formula, [\[1\]](#), [\[2\]](#). Formulas (1) and (2) employ finite differences, defined as follows:

$$f_{i+1/2}^1 = f_{i+1} - f_i \quad f_i^m = f_{i+1/2}^{m-1} - f_{i-1/2}^{m-1}.$$

The advantage of Gauss' interpolation formulas consists in the fact that this selection of interpolation nodes ensures the best approximation of the residual term of all possible choices, while the ordering of the nodes by their distances from the interpolation point reduces the numerical error in the interpolation.

## Stirling Interploation

Stirling Approximation or Stirling Interpolation Formula is an interpolation technique, which is used to obtain the value of a function at an intermediate point within the range of a discrete set of known data points .

Stirling Formula is obtained by taking the average or mean of the Gauss Forward and Gauss Backward Formula . Both the Gauss Forward and Backward formula are formulas for obtaining the value of the function near the middle of the tabulated set .

### How to find

Stirling Approximation involves the use of forward difference table, which can be prepared from the given set of x and f(x) or y as given below –

x	y	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
$x_0$	$y_0$					
$x_1$	$y_1$	$\Delta y_0$				
$x_2$	$y_2$	$\Delta y_1$	$\Delta^2 y_0$			
$x_3$	$y_3$	$\Delta y_2$	$\Delta^2 y_1$	$\Delta^3 y_0$		
$x_4$	$y_4$	$\Delta y_3$	$\Delta^2 y_2$	$\Delta^3 y_1$	$\Delta^4 y_0$	$\Delta^5 y_0$
$x_5$	$y_5$	$\Delta y_4$	$\Delta^2 y_3$	$\Delta^3 y_2$	$\Delta^4 y_1$	

This table is prepared with the help of x and its corresponding f(x) or y . Then, each of the next column values is computed by calculating the difference between its preceding and succeeding values in the previous column, like

$$\Delta y_0 = y_1 - y_0, \Delta y_1 = y_2 - y_1,$$

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0, \text{ and so on.}$$

Now, the Gauss Forward Formula for obtaining f(x) or y at **a** is:

where,  $p = \frac{a - x_0}{h}$  ,

**a** is the point where we have to determine f(x),  $x$  is the selected value from the given x which is closer to **a** (generally, a value from the middle of the table is selected), and h is the difference between any two consecutive x. Now,  $y$  becomes the value corresponding to  $x$  and values before  $x$  have negative subscript and those after have positive subscript, as shown in the table below –

x	y	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
$\vdots$	$\vdots$				
$x_{-2}$	$y_{-2}$	$\Delta y_{-2}$			
$x_{-1}$	$y_{-1}$	$\Delta y_{-1}$	$\Delta^2 y_{-2}$		
$x_0$	$y_0$		$\Delta^2 y_{-1}$	$\Delta^3 y_{-2}$	
$x_1$	$y_1$	$\Delta y_0$	$\Delta^2 y_0$	$\Delta^3 y_{-1}$	$\Delta^4 y_{-2}$
$x_2$	$y_2$	$\Delta y_1$			
$\vdots$	$\vdots$				

Stirling's Formula gives a good approximation for  $n!$  in terms of elementary functions. Before stating the formula, we introduce the following notation:

if  $f(n)$  is a function and  $g(n)$  is a function, then we write  $f(n) \sim g(n) \leftrightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1$ . The statement  $f(n) \sim g(n)$  is read  $f(n)$  is asymptotic to  $g(n)$  as  $n \rightarrow \infty$ .

For example, one verifies that  $n^2 \sim (n+1)^2$  and  $\sqrt{1+n} \sim \sqrt{n}$ .

Here is Stirling's Formula: Stirling's Formula  $n! \sim nn e^{-n} \sqrt{2\pi n}$ . The following graph shows a plot of the function  $h(n) = n!/nn e^{-n} \sqrt{2\pi n}$ , confirming Stirling's Formula:  $h(n) \rightarrow 1$  as  $n \rightarrow \infty$ . It turns out that  $h(n)$  is decreasing so  $nn e^{-n} \sqrt{2\pi n}$  always underestimates  $n!$  by a small amount.

10 20 30 40 50 60 70 1.03 1.04 1.02 1.01 Figure 1 :

Stirling's Formula The proof of Stirling's Formula is beyond the scope of this course. Instead of proving the formula, we rather give a proof of a weaker statement: we show that for every positive integer  $n$ ,  $nn e^{-n} < n! < (n+1)n+1e^{-n}$ . (1) This does not prove Stirling's Formula, but it gives motivation for the  $nn e^{-n}$  term in the formula. The proof of the  $\sqrt{2\pi n}$  part of the formula is more difficult. 1 First Proof — To prove (1), we just have to show (by taking logarithms):  $n \log n - n < \log(n!) < (n+1) \log(n+1) - n$ . Since  $n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$ ,  $\log(n!) = \log 1 + \log 2 + \dots + \log n$ . The sum on the right can be estimated by integrals: let's show that  $\log 1 + \log 2 + \dots + \log n < \int_1^{n+1} \log x \, dx$ . To see this, note that the integral represents the area under the curve  $y = \log x$  (the red curve in the left plot below) for  $1 \leq x \leq n+1$ , whereas the sum  $\log 1 + \log 2 + \dots + \log n$  represents adding up the areas of rectangles with height  $\log k$  for  $k = 1, 2, \dots, n$  (see green step function in the left plot below). Now we can work out the integral:  $\int_1^{n+1} \log x \, dx = x \log x - x + 1 \Big|_1^{n+1} = (n+1) \log(n+1) - n$ . Therefore  $\log(n!) < (n+1) \log(n+1) - n$ . We're going to do the same thing to prove  $\log(n!) > n \log n - n$ : we claim that  $\log 1 + \log 2 + \dots + \log n > \int_0^n \log x \, dx$ .

This is shown in the figure on the right, with the red curve representing  $\log x$  and the rectangles representing  $\log(1) + \log(2) + \dots + \log(n)$ . 2.0 0.5 x 2 6 1.5 1.0 0.0 4 8 2 x 1.5 0.5 2.0 4 6 0.0 1.0 8 Figure 2 : Approximating  $\log(n!)$  Therefore  $\log(n!) > \int_0^n \log x \, dx = n \log n - n$ , which completes the proof of (1).



## Bessel's Interpolation

**Interpolation** is the technique of estimating the value of a function for any intermediate value of the independent variable, while the process of computing the value of the function outside the given range is called **extrapolation**.

**Central differences** : The central difference operator  $\delta$  is defined by the relations :

$$y_1 - y_0 = \delta y_{1/2}, y_2 - y_1 = \delta y_{3/2}, \dots, y_n - y_{n-1} = \delta y_{n - \frac{1}{2}}.$$

Similarly, high order central differences are defined as :

$$\delta y_{3/2} - \delta y_{1/2} = \delta^2 y_1, \quad \delta y_{5/2} - \delta y_{3/2} = \delta^2 y_2$$

**Note** – The central differences on the same horizontal line have the same suffix

**Central difference table**

$x$	$y$	$\delta y$	$\delta^2 y$	$\delta^3 y$	$\delta^4 y$	$\delta^5 y$
$x_0$	$y_0$					
$x_1$	$y_1$	$\delta y_{1/2}$				
$x_2$	$y_2$	$\delta y_{3/2}$	$\delta^2 y_1$			
$x_3$	$y_3$	$\delta y_{5/2}$	$\delta^2 y_2$	$\delta^3 y_{3/2}$		
$x_4$	$y_4$	$\delta y_{7/2}$	$\delta^2 y_3$	$\delta^3 y_{5/2}$	$\delta^4 y_2$	
$x_5$	$y_5$	$\delta y_{9/2}$	$\delta^2 y_4$	$\delta^3 y_{7/2}$	$\delta^4 y_3$	$\delta^5 y_{5/2}$

**Bessel's Interpolation formula** –

$$\begin{aligned}
 f(u) = & \left\{ \frac{f(0) + f(1)}{2} \right\} + \left( u - \frac{1}{2} \right) \Delta f(0) \\
 & + \frac{u(u-1)}{2!} \left\{ \frac{\Delta^2 f(-1) + \Delta^2 f(0)}{2} \right\} \\
 & + \frac{(u-1) \left( u - \frac{1}{2} \right) u}{3!} \Delta^3 f(-1) \\
 & + \frac{(u+1) u (u-1) (u-2)}{4!} \left\{ \frac{\Delta^4 f(-2) + \Delta^4 f(-1)}{2} \right\} + \dots
 \end{aligned}$$

It is very useful when  $u = 1/2$ . It gives a better estimate when  $1/4 < u < 3/4$ . Here  $f(0)$  is the origin point usually taken to be mid point, since Bessel's is used to interpolate near the centre.  $h$  is called the interval of difference and  $u = (x - f(0)) / h$ , Here  $f(0)$  is term at the origin chosen.

**Examples –**

Input	:	Value	at	27.4	?	
$x$ :	25	26	27	28	29	30
$f(x)$ :	4.000	3.846	3.704	3.571	3.448	3.333.

**Output**

$u$	$10^3 f(u)$	$10^3 \Delta f(u)$	$10^3 \Delta^2 f(u)$	$10^3 \Delta^3 f(u)$	$10^3 \Delta^4 f(u)$	$10^3 \Delta^5 f(u)$
-2	4000	-154				
-1	3847	-142	12			
→ 0	3704	-133	9	-3	4	
1	3571	-123	10	1	-3	-7
→ 2	3448	-115	8	-2		
3	3333					

Value at 27.4 is 3.64968